

604639

604639

1
G
U

THE OPTIMIZATION OF A QUADRATIC FUNCTION
SUBJECT TO LINEAR CONSTRAINTS

Harry Markowitz

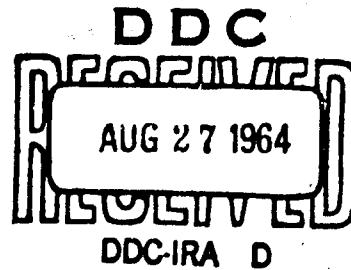
P-637

Revised 27 June 1955

Approved for OTS release

35¢

COPY	1 OF 1	fp
HARD COPY	\$.20	
MICROFICHE	\$.050	



THE OPTIMIZATION OF A QUADRATIC FUNCTION

SUBJECT TO LINEAR CONSTRAINTS

Harry Markowitz*

1. Quadratic Problems

Suppose that variables X_1, \dots, X_N are to be chosen subject to linear constraints:

- 1) $\sum a_{ij} X_j = b_i \quad i = 1, \dots, m_1$
- 2) $\sum a_{ij} X_j \geq b_i \quad i = m_1 + 1, \dots, m$
- 3) $X_j \geq 0 \quad j = 1, \dots, N_1$

where $0 \leq m_1 \leq m$, $0 \leq N_1 \leq N$ and the matrix (a_{ij}) $i = 1, \dots, m_1$ has rank m_1 (otherwise the system is inconsistent or has at least one redundant equation). The payoff is a linear function $R = \sum r_j X_j$ whose coefficients r_j are not known at the time the X_j are chosen. The r_j , rather, are random variables with expected values μ_j and covariances σ_{jk} (including variances $\sigma_{jj} = \sigma_j^2$). The expected value of R is

$$4) \quad E = \sum \mu_j X_j$$

The variance of R is

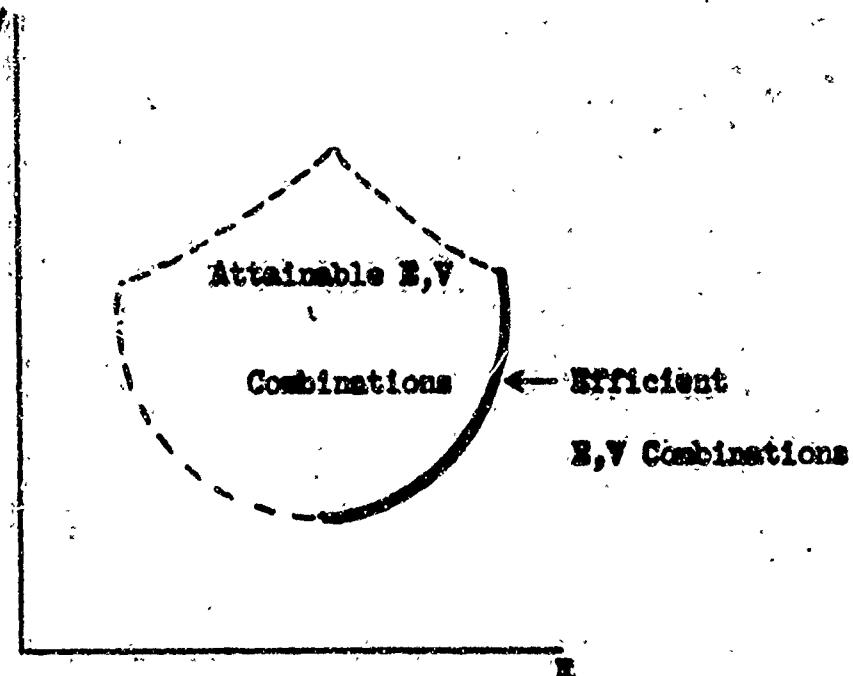
$$5) \quad V = \sum \sigma_{jk} X_j X_k$$

Suppose further that some decision-maker likes expected payoff (E) and dislikes variance of payoff (V). Our problem is to compute for the decision-maker (a) the "efficient combinations" of E and V , i.e., those

* The writer has particularly benefited from discussions with Kenneth Arrow on the subject matter of this paper.

attainable (E, V) combinations which give minimum V for given E and maximum E for given V (see Figure 1); and (b) the points in the X space associated with the efficient E, V combinations, i.e., the set of efficient X 's.

Figure 1.



A computing technique is presented in this paper for generating the above efficient sets. An adaptation of this technique can be used for problems of maximizing or minimizing quadratic forms (with the "right" properties) subject to linear constraints.

The practical problem which first suggested the above computing problem was that of selecting a portfolio of securities.* Here X_j is the amount invested in the j^{th} security; the μ_j and c_{jk} are the expected returns and covariances of returns from the various securities. In the simplest case the constraint set is $\sum X_j = 1$, $X_j \geq 0$. A problem of very similar structure analyzed independently by H. S. Houthakker,**

*Harry Markowitz, "Portfolio Selection," Journal of Finance 1952.

**"La Forme Des Courbes D'Engel," Officiers du Séminaire d'Econometrie 1953.

is that of finding the expenditure on various goods as a function of income for an individual whose utility function is of the form $u = \sum a_j X_j + \sum a_{ij} X_i X_j$. A problem of maximizing a monopolist's quadratic profit function subject to linear constraints is presented by Robert Dorfman.^{*} Another problem of this general character is that of maximizing a quadratic likelihood function where there is a priori information concerning the values of parameters to be estimated. Now that reasonably convenient computing procedures exist for such quadratic problems we may be permitted the hope that still other classes of interesting questions can be reduced to this form.

This paper will discuss only minimization problems involving the quadratic form $\sum a_{ij} X_i X_j$ whose matrix (a_{ij}) is positive semi-definite. The reader should have no difficulty in extending the results to minimization problems involving $\sum a_j X_j + \sum a_{ij} X_i X_j$ where (a_{ij}) is positive semi-definite or maximization problems where (a_{ij}) is negative semi-definite.

2. Assumptions

According to customary usage we say:

a) A set of points (S) (In Euclidean n-space) is convex if $x^{(1)} \in S$ and $x^{(2)} \in S$ imply $\lambda x^{(1)} + (1-\lambda) x^{(2)} \in S$, for any $0 \leq \lambda \leq 1$.

b) A set is closed if $x_1, \dots, x_t, \dots \rightarrow y$ and $x_1, \dots, x_t, \dots \in S$ imply $y \in S$.

c) A function $f(x)$ is convex over a set S if $x^{(1)} \in S$, $x^{(2)} \in S$ and $\lambda x^{(1)} + (1-\lambda) x^{(2)} \in S$ imply $f(\lambda x^{(1)} + (1-\lambda) x^{(2)}) \leq \lambda f(x^{(1)}) + (1-\lambda) f(x^{(2)})$ for all $0 \leq \lambda \leq 1$.

^{*}Application of Linear Programming to the Theory of the Firm, University of California Press, 1951.

d) A function is strictly convex over a set S if $x^{(1)} \in S$, $x^{(2)} \in S$ and $\lambda x^{(1)} + (1-\lambda) x^{(2)} \in S$ imply $f(\lambda x^{(1)} + (1-\lambda) x^{(2)}) < \lambda f(x^{(1)}) + (1-\lambda) f(x^{(2)})$ for all $0 < \lambda < 1$.

The set of points \tilde{S} which satisfy constraints 1), 2), and 3) is a closed, convex set. Variance (V) is a positive semi-definite quadratic form, i.e., $\sum_j \sum_k c_{jk} X_j X_k \geq 0$ for all (X_1, \dots, X_p) . It is also convex.

The covariance matrix (c_{jk}) is non-singular if and only if V is positive definite $\left\{ \text{i.e., } \sum_j \sum_k c_{jk} X_j X_k > 0, \text{ of } (X_1, \dots, X_p) \neq (0, \dots, 0) \right\}$ which in turn is true if and only if V is strictly convex over the set of all X .

That $V(X) \geq 0$, for all X , is due to the fact that V is the expected value of a square, $E(r-Er)^2$, and therefore cannot be negative. That $|c_{ij}| \neq 0$ if and only if V is positive definite is a corollary of material found, e.g., in Birkhoff and MacLane, A Survey of Modern Algebra, Chapter IX, particularly section 8, pp. 243-247. The implications of positive definiteness and semi-definiteness for convexity may be seen as follows: Let $C = (c_{ij})$. Let X and Y be column vectors; X' and Y' be row vectors. C is symmetric so that $C = C'$ and $X'CY = Y'CX$ for any X, Y . We wish to see the implications of

$$\begin{aligned} (1) \quad & X'CX \geq 0 \\ (2) \quad & X'CX = 0 \iff X = 0 \end{aligned}$$

for the difference

$$D = \left\{ \lambda X'CX + (1-\lambda) Y'CY \right\} - \left\{ (\lambda X' + (1-\lambda) Y') C (\lambda X + (1-\lambda) Y) \right\}$$

Expanding the second term and subtracting we get

$$\begin{aligned} D &= \lambda(1-\lambda) \cdot \left\{ X'CX - 2X'CY + Y'CY \right\} \\ &= \lambda(1-\lambda) \cdot \left\{ (X'-Y')C(X-Y) \right\} \end{aligned}$$

Assumption (1) implies $D > 0$ for all X, Y . Assumptions (1) and (2) imply $D > 0$ if $X \neq Y$. Conversely if D is positive for all $X \neq Y$, letting $Y = 0$ we find $X'CX > 0$ for all $X \neq 0$.

We will assume that \tilde{S} is not vacuous. We will also assume that V is strictly convex over the set of X 's which satisfy the equations

$$\sum a_{ij} X_j = b_i \quad i = 1, \dots, m_1$$

This assures us that V takes on a unique minimum over \tilde{S} and that if $S = S^0$ is attainable in \tilde{S} , then V takes on unique minimum over the set

$$\tilde{S} \cap \left\{ X \mid \sum \mu_j X_j \geq E_0 \right\}$$

If a function is convex over a set S it is convex over any subset of S ;

therefore $|a_{ij}| \neq 0$ implies that V is strictly convex over $\left\{ X \mid \sum a_{ij} X_j = b_i, i = 1, \dots, m_1 \right\}$. This is not a necessary condition however. Necessary and sufficient conditions on A and $(a_{ij}) = C$ are discussed in the footnote.^{**}

^{*}Since equations 1) have rank m_1 , m_1 variables and the m_1 equations could be eliminated (as in footnote ^{**}) to leave a system with $N-m_1$ variables and $(m-m_1) + N_1$ inequalities. V is strictly convex in these $N-m_1$ variables and therefore the associated quadratic is positive definite. Let Y be any point in the space of the $N-m_1$ variables satisfying the $(m-m_1) + N_1$ inequalities. The points which satisfy these inequalities and have $V \leq V(Y)$ form a compact, convex set. Since V is continuous it attains its minimum at least once on this set. Since it is strictly convex it attains its minimum only once. The same argument applies if the constraint

$$\sum \mu_j X_j \geq E^0$$

is added to the $(m-m_1) + N_1$ inequalities.

^{**}Suppose $m_1 \geq 1$. Since $\begin{pmatrix} a_{11} & \dots & a_{1N} \\ \vdots & \ddots & \vdots \\ a_{m_1 1} & \dots & a_{m_1 N} \end{pmatrix}$ has rank m_1 we can write,

after perhaps relabeling variables,

$$\begin{pmatrix} a_{11} & \dots & a_{1m_1} \\ \vdots & \ddots & \vdots \\ a_{m_1 1} & \dots & a_{m_1 m_1} \end{pmatrix} \begin{pmatrix} X_1 \\ \vdots \\ X_{m_1} \end{pmatrix} + \begin{pmatrix} a_{1m_1+1} & \dots & a_{1N} \\ \vdots & \ddots & \vdots \\ a_{m_1 m_1+1} & \dots & a_{m_1 N} \end{pmatrix} \begin{pmatrix} X_{m_1+1} \\ \vdots \\ X_N \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_{m_1} \end{pmatrix}$$

or $A^{(1)} X^{(1)} + A^{(2)} X^{(2)} = b$ where $A^{(1)}$ is non singular. We thus have

3. The Critical Line \bar{x}

Let us first note the answer to a simpler problem than that of finding all (\bar{z}, \bar{v}) efficient points. Suppose we wished to minimize V subject to the equations

$$\sum c_{ij} x_j = b_i \quad i = 1, \dots, n_1$$

without regard to the inequalities 2 and 3. A necessary condition for a minimum is that x_1, \dots, x_n be a solution to the Lagrangian equations,

$$6) \frac{\partial \sum c_{ij} x_j - 2 \sum \lambda_k \sum c_{ij} x_j}{\partial x_k} = 0$$

$$k = 1, \dots, n$$

$\bar{x}^{(1)} = (A^{(1)})^{-1} b - (A^{(1)})^{-1} A^{(2)} \bar{x}^{(2)}$. We can express V in terms of $\bar{x}^{(2)}$ by substitution, i.e.,

$$\bar{x} = \begin{pmatrix} \bar{x}^{(1)} \\ \bar{x}^{(2)} \end{pmatrix} = \begin{pmatrix} A^{(1)-1} b \\ 0 \end{pmatrix} = \begin{pmatrix} A^{(1)-1} & A^{(2)} \\ 0 & I \end{pmatrix} \bar{x}^{(2)}$$

$$V = \bar{x}' C \bar{x} = (\bar{x}^{(1)'} \bar{x}^{(2)'}) C \begin{pmatrix} \bar{x}^{(1)} \\ \bar{x}^{(2)} \end{pmatrix}$$

$$= [b' A^{(1)-1} b, 0] - (\bar{x}^{(2)'} A^{(2)'} A^{(1)-1} A^{(2)}, \bar{x}^{(2)'}) C \left[\begin{pmatrix} A^{(1)-1} b \\ 0 \end{pmatrix} - \begin{pmatrix} A^{(1)-1} A^{(2)} \bar{x}^{(2)} \\ \bar{x}^{(2)} \end{pmatrix} \right]$$

$$= (b' A^{(1)-1} b, 0) C \begin{pmatrix} A^{(1)-1} b \\ 0 \end{pmatrix} - 2 \bar{x}^{(2)'} (A^{(2)'} A^{(1)-1}, I) C \begin{pmatrix} A^{(1)-1} b \\ 0 \end{pmatrix}$$

$$+ \bar{x}^{(2)'} \left\{ (A^{(2)'} A^{(1)-1}, I) C \begin{pmatrix} A^{(1)-1} A^{(2)} \\ I \end{pmatrix} \right\} \bar{x}^{(2)}$$

V is strictly convex for all $\bar{x}^{(2)}$ if and only if the last (i.e., the quadratic) term is strictly convex. This is so if and only if

$$(A^{(2)'} A^{(2)-1}, I) C \begin{pmatrix} A^{(1)-1} A^{(2)} \\ I \end{pmatrix}$$

is non-singular.

as well as $\sum a_{ij} x_j = b_i \quad i = 1, \dots, n_1$

i.e.,

$$7) \sum_{j=1}^N a_{kj} x_j + \sum (-\lambda_i) a_{ik} = 0 \quad k = 1, \dots, n_1$$

$$8) \sum a_{ij} x_j = b_i \quad i = 1, \dots, n_1$$

Given the assumption that V is strictly convex over $\left\{ \mathbf{x} \mid \sum a_{ij} x_j = b_i, i = 1, \dots, n_1 \right\}$ it follows that

$$\begin{pmatrix} a_{11} & \cdots & a_{1N} & a_{11} & \cdots & a_{n_1 1} \\ \vdots & \ddots & \vdots & \vdots & & \vdots \\ a_{N1} & \cdots & a_{NN} & a_{N1} & \cdots & a_{n_1 N} \\ a_{11} & \cdots & a_{1N} & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{n_1 1} & \cdots & a_{n_1 N} & 0 & \cdots & 0 \end{pmatrix}$$

is non-singular.* Since a strictly convex V takes on a unique minimum on a convex set, the unique solution to 7) and 8) is this minimum.

Next consider the problem of minimizing V subject not only to 1) but also to the constraint that $\mathbf{x} = \mathbf{B}_0$, i.e.,

* If $n_1 = 0$ the statement reduces to one proved in the footnote of p. 4. Suppose $n_1 \geq 1$. If $\begin{pmatrix} C & A \\ A & 0 \end{pmatrix}$ is singular there is a vector $\begin{pmatrix} Y \\ -\lambda \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ such that $\begin{pmatrix} C & A \\ A & 0 \end{pmatrix} \begin{pmatrix} Y \\ -\lambda \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ i.e., $\begin{pmatrix} CY \\ AY \end{pmatrix} = \begin{pmatrix} A'\lambda \\ 0 \end{pmatrix}$

Since the rank of A is n_1 there is no $\lambda \neq 0$ such that $A'\lambda = 0$, therefore $Y \neq 0$ in $\begin{pmatrix} Y \\ -\lambda \end{pmatrix}$ above. $V(Y) = Y'C Y = Y'A'\lambda = (AY)'\lambda = 0$
Let X be any point in $S = \{ \mathbf{x} \mid AX = b \}$ where $\mathbf{x}' = (x_1, \dots, x_N)$

$$9) \sum_{j=1}^N \mu_j x_j = E_0$$

We must distinguish two cases:

Case 1: The row vector (μ_1, \dots, μ_N) can be expressed as a linear combination of the (a_{11}, \dots, a_{1N}) , i.e., there exists $\alpha_1, \dots, \alpha_{m_1}$ such that

$$10) (\alpha_1, \dots, \alpha_{m_1}) \begin{pmatrix} a_{11} & \cdots & a_{1N} \\ \vdots & \ddots & \vdots \\ a_{m_1 1} & \cdots & a_{m_1 N} \end{pmatrix} = (\mu_1, \dots, \mu_N)$$

Case 2: There does not exist such a linear combination.

As is shown below,* in Case 1 only one value of E , say $E = E^*$, is attainable. Therefore if we require $E \neq E^*$ no solution can be found; if we require $E = E^*$ equations, 7) and 8) give the minimum. In Case 2 the matrix

In. cont'd. from p. 7

$$b' = (b_1, \dots, b_{m_1})$$

$A(X+Y) = AX + 0 = b$; therefore $X+Y$ is in S

$$V(X) = X'CX$$

$$V(X+Y) = X'CX + 2X'CY = X'CX + 2X'A'\lambda = X'CX + 2b'\lambda$$

$$V(1/2X + 1/2(X+Y)) = V(X + 1/2Y)$$

$$= X'CX + X'CY$$

$$= X'CX + b'\lambda$$

$$= 1/2 V(X) + 1/2 V(X+Y)$$

Thus contradicting strict convexity.

*If $\alpha'A = \mu'$ and $AX = b$ then $E = \mu'X = \alpha'AX = \alpha'b$

$$\begin{array}{cccccc} \sigma_{11} & \dots & \sigma_{1N} & a_{11} & \dots & a_{m_1 N} & \mu_1 \\ \vdots & & \vdots & \vdots & & \vdots & \vdots \\ \sigma_{N1} & \dots & \sigma_{NN} & a_{1N} & \dots & a_{m_1 N} & \mu_N \\ \\ a_{11} & \dots & a_{m_1 1} & 0 & \dots & 0 & 0 \\ \vdots & & \vdots & \vdots & & \vdots & \vdots \\ a_{m_1 1} & \dots & a_{m_1 N} & 0 & \dots & 0 & 0 \\ \\ \mu_1 & \dots & \mu_N & 0 & \dots & 0 & 0 \end{array}$$

is non-singular* and therefore the equations

$$11) \quad \sum \sigma_{jk} x_k + \sum (-\lambda_i) a_{ij} - \lambda_B \mu_j = 0 \quad j = 1, \dots, N$$

$$12) \quad \sum a_{ij} x_j = b_i \quad i = 1, \dots, m_1$$

$$13) \quad \sum \mu_j x_j = E^0$$

have a unique solution which gives minimum V for the specified E^0 . If we let E^0 go from $-\infty$ to $+\infty$ the solution to 11), 12), and 13) traces out a line in the (X, λ) space. This line may also be described as the solution to the following $N + m_1$ equations in $N + m_1 + 1$ unknowns:

$$14) \quad \sum_{j=1}^N \sigma_{jk} x_k - \sum_{i=1}^{m_1} \lambda_i a_{ij} - \lambda_B \mu_j = 0 \quad j = 1, \dots, N$$

$$15) \quad \sum a_{ij} x_j = b_i \quad i = 1, \dots, m_1$$

or

$$16) \quad \sum \sigma_{jk} x_k + \sum (-\lambda_j) a_{ij} = \lambda_B \mu_j \quad j = 1, \dots, N$$

$$17) \quad \sum a_{ij} x_j = b_i \quad i = 1, \dots, m_1$$

for $-\infty \leq \lambda_B \leq +\infty$

*Same proof as in the footnote on page 7 using the fact that

$$\begin{pmatrix} A \\ \mu \end{pmatrix} \quad \text{has rank } m_1 + 1.$$

Since the matrix of equations 16) and 17) is non-singular, given our assumption of strict convexity, they have a solution for every value of λ_B whether we are in Case 1 or Case 2 above. In Case 1 the values of X_1, \dots, X_N do not change (only the values of the λ 's change) as λ_B goes from $-\infty$ to $+\infty$.^{*} In Case 2 the X 's as well as the λ 's change. In Case 2 we can define $\hat{V}(E)$ to be minimum V as a function of E . $D_E = \frac{d\hat{V}}{dE}$. $\hat{V}(E)$ must be strictly convex; therefore E increases with λ_B . In Section 11 we show that $\hat{V}(E)$ is a parabola.

4. Critical Lines $\mathcal{A}(\mathcal{I}, \mathcal{J})$

The set of points (X, λ) which satisfy 16) and 17) will be referred to as the critical line $\bar{\mathcal{S}}$ associated with the subspace.

$$\bar{\mathcal{S}} = \left\{ X \mid \sum a_{ij} X_j = b_i \text{ for } i = 1, \dots, m_1 \right\}$$

Critical lines will also be associated with certain other subspaces.

Let X_{j_1}, \dots, X_{j_J} be a subset of variables. Let

$$\sum a_{ij} X_j = b_i \quad i = i_1, \dots, i_I$$

be a subset of the constraints 1) and 2) with the inequalities replaced by equalities when $i > m_1$. Let \mathcal{I} be the ordered set of indices $\{i_1, \dots, i_I\}$; let \mathcal{J} be the ordered set $\{j_1, \dots, j_J\}$. We will be particularly interested in \mathcal{I} and \mathcal{J} of the form

$$18) \quad \mathcal{I} = \{1, 2, \dots, m_1, i_{m_1+1}, \dots, i_I\} \quad \text{where } I \geq m_1$$

$$19) \quad \mathcal{J} = \{j_1, \dots, j_L, N_1 + 1, \dots, N_1 + L\} \quad \text{where } 0 \leq L \leq N_1$$

*For if $\mu = A'\lambda$ and $\begin{pmatrix} C & A' \\ A & 0 \end{pmatrix} \begin{pmatrix} X^0 \\ \lambda^0 \end{pmatrix} = R + \begin{pmatrix} \mu \\ 0 \end{pmatrix} \lambda_B^0$
then $\begin{pmatrix} C & A' \\ A & 0 \end{pmatrix} \begin{pmatrix} X^0 \\ \lambda_B^0 + \lambda^0 \end{pmatrix} = R + \begin{pmatrix} \mu \\ 0 \end{pmatrix} \begin{pmatrix} \lambda_B^0 + \epsilon \\ \lambda^0 \end{pmatrix}$

For any indices \mathcal{I} and \mathcal{J} satisfying 18) and 19) we define the submatrix

$$20) \quad A_{\mathcal{I}\mathcal{J}} = \begin{pmatrix} a_{i_1 j_1} & \cdots & a_{i_1 j_J} \\ \vdots & \ddots & \vdots \\ \vdots & \ddots & \vdots \\ a_{i_T j_1} & \cdots & a_{i_T j_J} \end{pmatrix}$$

We similarly define subvectors $x_{\mathcal{J}}$, $\lambda_{\mathcal{I}}$ and

$$21) \quad b_{\mathcal{I}} = \begin{pmatrix} b_{i_1} \\ \vdots \\ b_{i_T} \end{pmatrix},$$

also submatrices

$$22) \quad C_{\mathcal{J}\mathcal{J}} = \begin{pmatrix} c_{j_1 j_1} & \cdots & c_{j_1 j_J} \\ \vdots & \ddots & \vdots \\ \vdots & \ddots & \vdots \\ c_{j_J j_1} & \cdots & c_{j_J j_J} \end{pmatrix}$$

$$23) \quad M_{\mathcal{I}\mathcal{J}} = \begin{pmatrix} C_{\mathcal{J}\mathcal{J}} & A'_{\mathcal{I}\mathcal{J}} \\ A_{\mathcal{I}\mathcal{J}} & 0 \end{pmatrix}$$

If $I = m_I = 0$, \mathcal{I} is empty. In this case it will sometimes be convenient to think of $A_{\mathcal{I}\mathcal{J}}$ as having no rows and J columns. To every $(\mathcal{I}, \mathcal{J})$ satisfying 18) and 19) we associate a subspace

$$S(\mathcal{I}, \mathcal{J}) = \{x \mid x_j = 0 \text{ for } j \notin \mathcal{J}, A_{\mathcal{I}\mathcal{J}} x_{\mathcal{J}} = b_{\mathcal{I}}\}$$

If $A_{\mathcal{I}\mathcal{J}}$ has no rows this reduces to $S(\mathcal{J}) = \{x \mid x_j = 0 \text{ for } j \notin \mathcal{J}\}$

Since \mathcal{I} and \mathcal{J} satisfy 18) and 19) $S(\mathcal{I}, \mathcal{J}) \subset \bar{S}$. Since V is strictly convex over \bar{S} it is strictly convex over $S(\mathcal{I}, \mathcal{J})$.

$A_{\mathcal{J}f}$ has a rank of I or less. If $A_{\mathcal{J}f}$ has rank I then the matrix

$$24) \quad M_{\mathcal{J}f} = \begin{pmatrix} C_{\mathcal{J}f} & A_{\mathcal{J}f} \\ A_{\mathcal{J}f} & 0 \end{pmatrix}$$

is non-singular. (This is a special case of the proposition proved in the footnote on page 7.) If $A_{\mathcal{J}f}$ has rank less than I, $M_{\mathcal{J}f}$ is singular, for its last I rows are not independent. For every (\mathcal{J}, f) satisfying 18) and 19) whose $A_{\mathcal{J}f}$ has rank equal to the number of its rows, we define the critical line $\mathcal{L}_{\mathcal{J}f}$ to be the set of points $(x_1, \dots, x_n, \lambda_1, \dots, \lambda_n)$ which satisfy

$$25) \quad x_j = 0 \quad \text{for } j \notin \mathcal{J}$$

$$\lambda_i = 0 \quad \text{for } i \notin \mathcal{J}$$

and

$$\begin{pmatrix} x_j \\ -\lambda_i \end{pmatrix} = M_{\mathcal{J}f}^{-1} \begin{pmatrix} 0 \\ B_{\mathcal{J}} \end{pmatrix} + M_{\mathcal{J}f}^{-1} \begin{pmatrix} u_f \\ 0 \end{pmatrix} \quad \lambda_i$$

Equations 25) may be written in the form

$$26) \quad x_j = \alpha_{xj} + \beta_{xj} \lambda_B \quad \left. \right\} \quad -\infty < \lambda_B < \infty$$

$$27) \quad \lambda_i = \alpha_{\lambda i} + \beta_{\lambda i} \lambda_B$$

Equations 26) by themselves are the projection of $\mathcal{L}(\mathcal{J}, f)$ onto the X-space. As with \bar{S} and $\bar{\lambda}$ we have two cases:

1) Only one value of B is obtainable in $S(\mathcal{J}, f)$ and the X-projection is a point.

2) All values of B are obtainable and the X-projection is a line.

This line is the set of X's in $S(\mathcal{J}, f)$ which give minimum V for some B.

Let

$$28) \quad \xi_i = \sum a_{ij} x_j - b_i \quad i = 1, \dots, n$$

$$29) \quad \eta_j = 1/2 \frac{\partial V - 2 \sum_{i=1}^n \lambda_i \sum a_{ij} x_j - 2 \lambda_B \sum \mu_j}{\partial x_j}$$

$$= \sum_{j=1}^n c_{jk} x_k + \sum_i (-\lambda_i) a_{ij} - \mu_j \lambda_B$$

Constraints 1) and 2) state that

$$\xi_i = 0 \quad \text{for } i = 1, \dots, m_1$$

$$\xi_i \geq 0 \quad \text{for } i = m_1 + 1, \dots, n.$$

Along any critical line we have

$$30) \quad \begin{aligned} x_j &= 0 && \text{for } j \notin J \\ \eta_j &= 0 && \text{for } j \in J \\ \xi_i &= 0 && \text{for } i \in Q \\ \lambda_i &= 0 && \text{for } i \notin Q \end{aligned}$$

Also, from 25), letting m^{st} be the $(s,t)^{th}$ element of M_{gg}^{-1} we have

$$31) \quad x_{j_s} = \sum_{h=1}^J m^{s,h+j} b_{1h} + \left(\sum_{h=1}^J m^{sh} \cdot \mu_{j_h} \right) \lambda_B$$

$$= \alpha_{x_{j_s}} + \beta_{x_{j_s}} \lambda_B \quad \text{for } s = 1, \dots, J$$

$$32) \quad \lambda_{1s} = - \sum_{h=1}^I m^{s+j, h+j} b_{1h} - \left(\sum_{h=1}^J m^{s+j, h} \mu_{j_h} \right) \lambda_B$$

$$= \alpha_{\lambda_{1s}} + \beta_{\lambda_{1s}} \lambda_B \quad \text{for } s = 1, \dots, I$$

From 28) and 29) we have

$$33) \quad \xi_1 = \sum_{h=1}^J a_{1j_h} \alpha_{Xj_h} - b_1 + \left(\sum_{h=1}^J a_{1j_h} \beta_{Xj_h} \right) \lambda_B \\ = \alpha_{\xi_1} + \beta_{\xi_1} \lambda_B$$

$$34) \quad \eta_j = \left(\sum_{h=1}^J a_{jj_h} \alpha_{Xj_h} - \sum_{h=1}^I a_{ij_h} \alpha_{Xj_h} \right) \\ + \left(\sum_{h=1}^J a_{jj_h} \beta_{Xj_h} - \sum_{h=1}^I a_{ij_h} \beta_{Xj_h} - \mu_j \right) \lambda_B \\ = \alpha_{\eta_j} + \beta_{\eta_j} \lambda_B$$

A corollary of the results of an important paper by Kuhn and Tucker* is that a sufficient condition for a point \bar{X} to give minimum V for a set

$$\bar{S} \cap \left\{ X : \sum \mu_j X_j \geq z_0 \right\}$$

for some z_0 is that

$$35) \quad \begin{aligned} X_j &\geq 0 && \text{for } j \leq n_1 \text{ and in } J \\ \eta_j &\geq 0 && \text{for } j \notin J \\ \xi_1 &\geq 0 && \text{for } i \notin I \\ \lambda_i &\geq 0 && \text{for } i > n_1 \text{ and in } I \end{aligned}$$

and $\lambda_B \geq 0$. If $\lambda_B > 0$ the constraint is effective; if z_0 were increased an equally low value of V could not be obtained. If $\lambda_B = 0$, the point gives minimum V in \bar{S} . In either case the point is efficient.

It will be convenient at times to employ the following relabeling of variables:

* H. W. Kuhn and A. W. Tucker, "Nonlinear Programming" in Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability.

$$\begin{aligned}
 v_k &= x_k && \text{for } k = 1, \dots, N_1 \\
 36) \quad v_k &= \eta_{k-N_1} && \text{for } k = N_1 + 1, \dots, 2N_1 \\
 v_k &= \xi_{m_1+k - 2N_1} && \text{for } k = 2N_1 + 1, \dots, 2N_1 + m - m_1 \\
 v_k &= \lambda_{k + 2m_1 - m - 2N_1} && \text{for } k = 2N_1 + m - m_1 + 1, \dots, \\
 &&& 2N_1 + 2m - 2m_1
 \end{aligned}$$

Also

$$37) \quad K = 2N_1 + 2m - 2m_1$$

and

$$38) \quad \mathcal{K} = \left\{ \text{the set of } k \text{ which identify the variables in equation 30} \right\}$$

Thus on any critical line we have

$$39) \quad v_k = 0 \quad \text{for } k \in \mathcal{K}$$

and a point \bar{x} is efficient if it is a projection of a point on a critical line with

$$40) \quad v_k \geq 0 \quad \text{for } k \notin \mathcal{K}$$

or

$$41) \quad v_k \geq 0 \quad \text{for } k = 1, \dots, K$$

5. Intersections of Critical Lines; Non-Degeneracy Conditions

In the computing procedure of the next section we move along a critical line until it intersects a plane $v_k = 0$, $k = 1, \dots, K$. Then either one row and the corresponding column is added to M , or one row and the corresponding column is deleted from M . This raises two questions: (1) under what conditions will the matrix obtained by such additions or deletions be non-singular, and (2) how should the new

inverse be obtained? The latter question will not be discussed except to note that the possession of the old inverse is of great value in obtaining the new one.*

Concerning the former question, suppose $M_{\mathcal{J}\mathcal{J}}$, with \mathcal{J} satisfying 18) and 19), is non-singular and thus defines a critical line \mathcal{L} . Suppose \mathcal{L} intersects (but is not contained in) the plane $v_k = 0$, $1 \leq k \leq K$. We distinguish four cases depending on whether v corresponds to an X , an η , a λ or a ξ :

1. The deletion of a variable. Suppose \mathcal{L} intersects a plane $X_j = 0$, $j = 1, \dots, N_1$. Suppose that j is deleted from the set \mathcal{J} leaving the set \mathcal{J}^* . Is $M_{\mathcal{J}\mathcal{J}^*}$ non-singular? We may suppose without loss of generality that $j = J_1$ and that $A_{\mathcal{J}\mathcal{J}}$ may therefore be written

$$42) A_{\mathcal{J}\mathcal{J}} = (\alpha A_{\mathcal{J}\mathcal{J}^*})$$

where α is the column to be deleted. The matrix $\begin{pmatrix} \alpha & A_{\mathcal{J}\mathcal{J}^*} \\ 1 & 0 \dots 0 \end{pmatrix}$ has either rank, I or I + 1. If it has rank I, then

$$43) \left(\begin{array}{ccc|c} C_{\mathcal{J}\mathcal{J}} & A'_{\mathcal{J}\mathcal{J}} & \mu_j & \\ \vdots & \vdots & \vdots & \\ \mu_{j_1} & & & \\ \hline A_{\mathcal{J}\mathcal{J}} & 0 & 0 & \\ \vdots & \vdots & \vdots & \\ 1 & 0 \dots 0 & 0 \dots 0 & 0 \end{array} \right) = \tilde{M}$$

is singular. In this case the equations

* This involves procedures similar to those used in removing a variable from a regression analysis or modifying a basis in linear programming, e.g., see R. A. Fisher, Statistical Methods for Research Workers, p. 164, and R. Dorfman, Activity Analysis of Production and Allocation, 10th Ed., p. 358.

$$44) \quad \tilde{M} \begin{pmatrix} x_{j_1} \\ x_{j_2} \\ \vdots \\ x_{j_l} \\ \lambda_1 \\ \vdots \\ \lambda_m \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ b_1 \\ \vdots \\ b_{l-1} \\ I \\ 0 \end{pmatrix}$$

have either no solution or an infinity of solutions. Thus if 44) has one solution, i.e., if λ_i intersects $x_{j_1} = 0$ (but is not contained in it), \tilde{M} is non-singular and the rank of $\begin{pmatrix} \alpha & A_{gj_*}^* \\ 1 & 0 \end{pmatrix}$ is $I + 1$; hence the rank of $\begin{pmatrix} A_{gj_*}^* \\ 0 \dots 0 \end{pmatrix}$ is at least I . But the rank of $\begin{pmatrix} A_{gj_*}^* \\ 0 \dots 0 \end{pmatrix}$ = the rank of A_{gj_*} . Therefore the rank of A_{gj_*} is I and M_{gj_*} is non-singular.

2. The deletion of a constraint. Suppose λ_i intersects (but is not contained in) $x_{j_1} = 0$ for $i > m_1$. We may assume $i = i_1$ and that

$$45) \quad A_{gj_*} = \begin{pmatrix} A_{gj_*}^* \\ \alpha' \end{pmatrix}.$$

A_{gj_*} has rank I , $A_{gj_*}^*$ has rank $I-1$; therefore M_{gj_*} is non-singular.

3. Addition of a variable. Continuing the conventions used above, if A_{gj_*} has rank I , so has $A_{gj_*}^* = (A_{gj_*} \alpha)$. Therefore M_{gj_*} is non-singular.

4. Addition of a constraint. If λ_i intersects but is not contained in the plane $x_{j_1} = 0$, $i > m_1$ then

$$46) \quad \begin{pmatrix} C_{gj_*} & A_{gj_*} & u_{j_1} \\ A_{gj_*} & 0 & 0 \\ \alpha' & 0 & 0 \end{pmatrix} = \tilde{M}$$

(where α' is the row of coefficients of the new constraint) is non-singular.

Therefore $\begin{pmatrix} A_{ij} \\ \alpha \end{pmatrix} = A_{ij} * f$ has rank $I + 1$ and $M_{A_{ij} * f}$ is non-singular.

The tracing out of the efficient set is simplified if certain "accidents" do not occur. These accidents are described in the following "non-degeneracy" conditions. The next section of this paper presents a computing procedure for deriving the set of efficient points when all non-degeneracy conditions hold. In sections 7-10 these conditions are relaxed.

Condition 1. On no critical line do we have

$$v_k = \alpha_{vk} + \beta_{vk} \lambda_E \leq 0 \quad \text{for } k \notin \mathcal{K}$$

Condition 2. On any given critical line ℓ we do not have

$$\frac{-\alpha_{vk_1}}{\beta_{vk_1}} = \frac{-\alpha_{vk_2}}{\beta_{vk_2}} \quad \text{for any } k_1 \neq k_2 \text{ with } \beta_{vk_1} \neq 0,$$

$$\beta_{vk_2} \neq 0$$

Condition 3. E is bounded from above in \tilde{S} .

We will let L_g stand for "the linear programming problem of maximizing E subject to constraints 1), 2), and 3).

Condition 4. L_g has a unique non-degenerate solution.

Condition 4 implies condition 3.

6. The Algorithm under Conditions 1 through 4

We now assume that conditions 1 through 4 are satisfied. Condition 4 implies that the optimum solution to L_g has:

The following are corollaries of the basis and pricing theorems of linear programming. See, e.g., George B. Dantzig, Alex Orden, Philip Wolfe, "The Generalized Simplex Method for Minimizing a Linear Form under Linear Inequality Restraints," Pacific Journal of Mathematics, Vol. 5, No. 2, June, 1955.

- a) exactly m variables X_j and ξ_i are not at their lower extreme
- b) $A_{\varphi(1)} \varphi(1)$ (where $\varphi^{(1)}$ includes all i with $\xi_i = 0$ and $\varphi^{(1)}$ includes all j with X_j not at its lower limit) has rank equal to the number of its rows
- c) There exists "prices" p_i and "profitabilities" δ_j such that

$$47) p_i > 0 \quad \text{if } \xi_i = 0 \quad \text{for } i = n_1 + 1, \dots, m$$

$$48) p_i = 0 \quad \text{if } \xi_i > 0 \quad \text{for } i = n_1 + 1, \dots, m$$

$$49) \delta_j = \sum_i a_{ij} p_i + \mu_j$$

$$50) \delta_j = 0 \quad \text{for } j = n_1 + 1, \dots, N \quad \text{and for } X_j > 0 \quad j \leq n_1$$

$$51) \delta_j < 0 \quad \text{for } X_j = 0 \quad j \leq n_1$$

The matrix

$$52) M_{(1)} = \begin{pmatrix} C_{\varphi(1)} \varphi(1) & A'_{\varphi(1)} \varphi(1) \\ A_{\varphi(1)} \varphi(1) & 0 \end{pmatrix}$$

is non-singular and thus defines a critical line $\lambda_B^{(1)}$ along which

$$53) M_{(1)} \begin{pmatrix} x_{\varphi(1)} \\ -\lambda_B^{(1)} \varphi(1) \end{pmatrix} = \begin{pmatrix} 0 \\ B_{\varphi(1)} \end{pmatrix} + \begin{pmatrix} u_{\varphi(1)} \\ 0 \end{pmatrix} \cdot \lambda_B$$

Since $-A'_{\varphi(1)} \varphi(1) \cdot p_{\varphi(1)} = u_{\varphi(1)}$, if $x^0_{\varphi(1)} \lambda_B^0 \varphi(1)$ satisfy 53) for $\lambda_B = 0$, then

$$54) M_1 \begin{pmatrix} x^0_{\varphi(1)} \\ -\lambda_B^0 \varphi(1) - p_{\varphi(1)} \lambda_B \end{pmatrix} = \begin{pmatrix} 0 \\ B_{\varphi(1)} \end{pmatrix} + \begin{pmatrix} u_{\varphi(1)} \\ 0 \end{pmatrix} \lambda_B$$

for all λ_B .

Thus $\xi^{(1)}$ has

$$55) \quad \begin{aligned} x_j^{(1)} &= x_h^0 \\ \lambda_j^{(1)} &= \lambda_h^0 + p_j^{(1)} \lambda_E \end{aligned} \quad \left. \right\} \text{for all } \lambda_E$$

From 47) it follows that for sufficiently large λ_E

$$\lambda_i > 0 \quad \text{for } i > m_1 \text{ and in } \mathcal{J}$$

$$56) \quad \begin{aligned} \eta_j &= \sum c_{jh} x_h^0 - (\sum a_{ij} \lambda_i + \mu_j \lambda_E) \\ &= \sum c_{jh} x_h^0 - \sum a_{ij} \lambda_i^0 - (\sum a_{ij} p_i + \mu_j) \lambda_E \end{aligned}$$

51) implies that for sufficiently large λ_E $\eta_j > 0$ for $j \notin \mathcal{J}$. Thus for sufficiently large λ_E $\xi^{(1)}$ satisfies inequalities 40).

Let $\lambda_E^{(1)}$ be the largest value of λ_E at which $\xi^{(1)}$ intersects a plane $\eta_j = 0$ for $j \notin \mathcal{J}$ or $\lambda_i = 0$ for $i \in \mathcal{J}$. (The X and ξ do not vary along $\xi^{(1)}$.) If $\lambda_E^{(1)} \leq 0$ then X^0 gives minimum V as well as maximum E . Suppose $\lambda_E^{(1)} > 0$. Non-degeneracy condition 2 implies $\xi^{(1)}$ intersects only one plane $\eta_j = 0$ or $\lambda_i = 0$ at $\lambda_E^{(1)}$. In the former case we add j to \mathcal{J} ; in the latter case we delete i from \mathcal{J} , to form $\mathcal{J}^{(2)}$, $\mathcal{J}^{(2)}$.

The new matrix $M_{(2)} = M_{\mathcal{J}^{(2)}} \xi_{\mathcal{J}^{(2)}}$ is non-singular and defines a critical

line $\xi^{(2)}$. Suppose for a moment that it was $\eta_{j_0} = 0$ which $\xi^{(1)}$ intersected at $\lambda_E^{(1)}$. On $\xi^{(2)}$ we have at $\lambda_E = \lambda_E^{(1)}$:

$$57) \quad \begin{aligned} \lambda_i &> 0 \quad \text{for } i > m_1 \text{ and } i \in \mathcal{J}^{(2)} \\ \eta_j &> 0 \quad \text{for } j \notin \mathcal{J}^{(2)} \end{aligned}$$

$$\xi_i > 0 \quad \text{for } i \notin \mathcal{J}$$

$$\text{and } \begin{cases} x_{j_0} = 0 \\ x_j > 0 \quad \text{for all other } j \leq m_1 \text{ and } j \in \mathcal{J}^{(2)}. \end{cases}$$

As always $X_{j_0} = s + b\lambda_E^{(2)}$ along $\lambda_E^{(2)}$. Non-degeneracy condition 1 assures $b \neq 0$.

If $b < 0$ the projection of $\lambda_E^{(2)}$ would be efficient for $\lambda^* \geq \lambda_E^{(2)} \geq \lambda_E^{(1)}$ where $\lambda^* > \lambda_E^{(1)}$. This is impossible.* Therefore $b > 0$ and $\lambda_E^{(2)}$ is efficient for $\lambda_E^{(1)} \geq \lambda_E^{(2)} \geq \lambda_E^{(3)}$ where $\lambda_E^{(1)} > \lambda_E^{(2)}$. Similar remarks would apply if $\lambda_E^{(1)}$ first intersected $\lambda_{i_0} = 0$ and i_0 was deleted from \mathcal{J} .

$\lambda_E^{(2)}$ is the highest value of $\lambda_E < \lambda_E^{(1)}$ at which $\lambda_E^{(2)}$ intersects a plane $v_k = 0$ for $k = 1, \dots, K$. If this is an η_j we again add a j to \mathcal{J} . If it is a λ_i we delete i from \mathcal{I} ; if a ξ_i , we add i to \mathcal{I} ; if an X_i we delete j from \mathcal{J} . We form $M_{(3)}$ and $\lambda_{(3)}$ accordingly and find $\lambda_{(3)} < \lambda_{(2)}$. This process is repeated until $\lambda_E = 0$ is reached. At each step (s) $M_{(s)}$ is non-singular and if v_{k_s} is the new variable (η , X , λ , or ξ) which is no longer constrained to be zero we have at $\lambda^{(s-1)}$

$$58) \quad v_k > 0 \quad \text{for } k \neq k_s \text{ and } k \neq *$$

$$v_{k_s} = 0$$

By condition 1, $b_{v_{k_s}} \neq 0$ along $\lambda^{(s)}$. We argue below** that we cannot have $b_{v_{k_s}} < 0$. So $b_{v_{k_s}} > 0$ and $\lambda^{(s)}$ is efficient for $\lambda_E^{(s-1)} \geq \lambda_E \geq \lambda_E^{(s)}$ where $\lambda_E^{(s-1)} > \lambda_E^{(s)}$. Since there are only a finite number of critical

* Since $b \neq 0$ the X-projection of the critical line is a line rather than a point. Along this line E increases with $\lambda_E^{(1)}$. If $\lambda_E > \lambda_E^{(1)}$ were feasible then $E > E^{(1)} = \max E$ would be obtainable, which is impossible.

** If v_{k_s} is an X_j or ξ_i , $b_{v_{k_s}} < 0$ implies that there are two distinct points which minimize V for some $E > E^{(s-1)}$, which is impossible. This argument also applies if v_{k_s} is a λ_i or η_j unless the X-projection of the new critical line is a point. In the latter case we note (from the Kuhn and Tucker conditions) that an efficient point gives minimum $Q(\lambda_E) = V - \lambda_E E$ subject to 1), 2), and 3). For fixed λ_E , $Q(\lambda_E)$ has a unique minimum. If $v_{k_s} < 0$ then two distinct points give minimum $Q(\lambda_E)$ for some $\lambda_E > \lambda_E^{(s-1)}$.

lines, and each can satisfy inequalities 40) for only one segment $\lambda_k = 0$ is reached in a finite number of steps.

7. The Algorithm under Conditions 3 and 4

Let us now drop non-degeneracy conditions 1 and 2 but still assume conditions 3 and 4. We will use techniques analogous to the degeneracy-avoiding techniques of linear programming.*

For every number e we define a new problem $P(e)$ as follows:

$$\text{minimize } V(e) = \sum \sum c_{ij} x_i x_j + \sum e^j x_j$$

subject to

$$59) \quad \sum a_{ij} x_j = b_i + e^{N+i} \quad i = 1, \dots, n_1$$

$$60) \quad \sum a_{ij} x_j \geq b_i + e^{N+i} \quad i = n_1 + 1, \dots, m$$

$$61) \quad x_j \geq 0 \quad j = 1, \dots, N_2$$

For sufficiently small e the unique, optimal basis of L_g is feasible and, since it still satisfies the pricing conditions, is optimal.

* In linear programming these techniques are generally not needed in practice. In quadratic programming arbitrary selection of $v_k = 0$ with $b_{vk} < 0$ to go into X may (or may not) prove adequate. In any case the degeneracy handling techniques are available if needed. See George Dantzig, "Application of the Simplex Method to a Transportation Problem," Activity Analysis of Production and Allocation, Tjalling C. Koopmans, ed.; and A. Charnes, "Optimality and Degeneracy in Linear Programming," Econometrica, Vol. 20, No. 2, April, 1952, p. 160.

As we will see shortly for sufficiently small ϵ $P(\epsilon)$ satisfies non-degeneracy conditions 1) and 2). We will also see that for a sufficiently small ϵ^* the sequence of indices $(\ell_j)^s$ associated with the critical lines $\epsilon^{(s)}$, until $\lambda_{\underline{k}} = 0$ is reached, is the same for all $P(\epsilon)$ for $\epsilon^* \geq \epsilon > 0$. If we change indices (ℓ_j) in the same sequence as $P(\epsilon)$ for small ϵ , if we let $\lambda_{\underline{k}}$ decrease along any critical line when it can without violating $v_k \geq 0$, until we reach $\lambda_{\underline{k}} = 0$, then

- a) we will pass through a finite number of index sets each associated with a non-singular M_{ℓ_0} , before we reach $\lambda_{\underline{k}} = 0$.
- b) since $v_k \geq 0$ is maintained we have the desired solution to the original problem.

Along any critical line of $P(\epsilon)$ we have

$$(62) \begin{pmatrix} x_{ij} \\ -\lambda_{ij} \end{pmatrix} = K_{ij}^{-1} \begin{pmatrix} 0 \\ p_{ij} \end{pmatrix} + K_{ij}^{-1} \begin{pmatrix} \mu_{ij} \\ 0 \end{pmatrix} \cdot \lambda_B + K_{ij}^{-1} \begin{pmatrix} \lambda_1 \\ e \\ \vdots \\ \lambda_{I+j} \\ e \end{pmatrix}$$

or

$$(63) x_{ij_s} = \alpha_{x_{ij_s}} + \beta_{x_{ij_s}} \lambda_B + \sum_{h=1}^{I+j} m^{sh} e^{f(h)}$$

$$\text{where } f(1) = j_1, f(2) = j_2, \dots, f(I+j) = I + i_I$$

or

$$(64) x_{ij_s} = \alpha_{x_{ij_s}} + \beta_{x_{ij_s}} \lambda_B + p_{x_{ij_s}}(\epsilon)$$

Similarly

$$\begin{aligned} \lambda_{1_s} &= \alpha_{\lambda_{1_s}} + \beta_{\lambda_{1_s}} \lambda_B + \sum_{h=1}^{I+j} m^{j+h,h} e^{f(h)} \\ &= \alpha_{\lambda_{1_s}} + \beta_{\lambda_{1_s}} \lambda_B + p_{\lambda_{1_s}}(\epsilon) \end{aligned}$$

$$\begin{aligned} (65) \xi_1 &= \alpha_{\xi_1} + \beta_{\xi_1} \lambda_B + \sum_{s=1}^J a_{1,j_s} p_{x_{j_s}}(\epsilon) + e^{N+i} \\ &= \alpha_{\xi_1} + \beta_{\xi_1} \lambda_B + p_{\xi_1}(\epsilon) \end{aligned}$$

$$\begin{aligned} (66) \eta_j &= \alpha_{\eta_j} + \beta_{\eta_j} \lambda_B + \sum_{s=1}^J a_{j,j_s} p_{x_{j_s}}(\epsilon) \\ &\quad - \sum_{s=1}^I a_{j,s} p_{\lambda_{j_s}}(\epsilon) + e^j \\ &= \alpha_{\eta_j} + \beta_{\eta_j} \lambda_B + p_{\eta_j}(\epsilon) \end{aligned}$$

Consider the polynomials:

$$67) \quad p_{x_j}(e) \quad \text{for } j \in J$$

$$68) \quad p_{\lambda_i}(e) \quad \text{for } i \in I$$

$$69) \quad p_{v_i}(e) \quad \text{for } i \notin S$$

$$70) \quad p_{v_j}(e) \quad \text{for } j \notin J$$

None of the polynomials listed above have all zero coefficients, and no two have proportional coefficients. For: each polynomial of 69) and 70) has a term with a coefficient of 1 which every other polynomial has with a coefficient of zero. This leaves only the possibilities that some polynomial of 67) or 68) has all zero coefficients or two of these polynomials have proportional coefficients. Both these possibilities imply that M^{-1} is singular and are therefore impossible.

Since $p_{v_k}(e)$ has only a finite number of roots, for e sufficiently small.

$$p_{v_k}(e) \neq 0 \quad \text{for } k \notin K.$$

Thus

$$71) \quad v_k = \alpha_{vk} + \beta_{vk} \lambda_B + p_{vk}(e) \quad -\infty < \lambda_B < \infty$$

cannot be identically zero for $k \notin K$. The critical line intersects the plane $v_{k_1} = 0$ at

$$72) \quad \lambda_B' = \frac{-\alpha_{vk_1}}{\beta_{vk_1}} - \frac{p_{vk_1}(e)}{\beta_{vk_1}}$$

and the plane

$$v_{k_2} = 0 \text{ at}$$

$$73) \quad \lambda_E^* = \frac{-\alpha_{vk_2}}{\beta_{vk_2}} - \frac{p_{vk_2}(e)}{\beta_{vk_2}}$$

If, say,

$$74) \quad \frac{-\alpha_{vk_1}}{\beta_{vk_1}} > \frac{-\alpha_{vk_2}}{\beta_{vk_1}}$$

then for sufficiently small e

$$\lambda_E^* > \lambda_E''$$

On the other hand, since

$$p_{vk_1}(e) - p_{vk_2}(e) = 0 \quad k_1, k_2 \neq \infty$$

has a finite number of solutions, for sufficiently small e

$$\lambda_E^* \neq \lambda_E''$$

even if

$$75) \quad \frac{-\alpha_{vk_1}}{\beta_{vk_1}} = \frac{-\alpha_{vk_2}}{\beta_{vk_2}}$$

As $e \rightarrow 0$ the smallest power of e dominates; i.e., if, say, $\frac{-p_{vk_1}(e)}{\beta_{vk_1}}$

has an algebraically larger coefficient for the first power of e than $\lambda^* > \lambda''$ as $e \rightarrow 0$. If both have the same coefficient of e then it is the coefficients of e^2 that decide. And so on.

Since there are a finite number of critical lines and a finite number of planes $v_k = 0$ there is a single e^* such that for $e^* \geq e > 0$:

$P(e)$ satisfies non-degeneracy conditions 1) and 2); and the order of the index sets $(\mathcal{I}_k^j)^s$ is the same for all such e .

The n^{st} are needed for other purposes and are thus available for resolving degeneracy problems. The other coefficients of $p(e)$ can be

computed when needed.

8. The Algorithm when $L_{\bar{B}}$ is Degenerate but Unique

Suppose that the solution to $L_{\bar{B}}$ is degenerate in that one or more of the basis variables X_j or ξ_j is "accidentally" zero, but is unique in that $\delta_j < 0$ for all X_j not in the basis and $p_i > 0$ for all ξ_i not in the basis.

The constraints of $L_{\bar{B}}$ may be written as a system of equations including the ξ_i as variables:

$$76) \quad B \begin{pmatrix} X \\ \xi \end{pmatrix} = b$$

If \tilde{B} is the submatrix of optimal basis vectors and if X_j and ξ_j are the optimal basis variables then the optimal solution is given by

$$77) \quad \begin{pmatrix} X^0 \\ X^0 \\ \xi^0 \\ \xi^0 \end{pmatrix} = \tilde{B}^{-1} b$$

while all other variables are zero. After we solve $L_{\bar{B}}$ we may modify it, forming $L_{\bar{B}}(e)$ as follows

$$78) \quad B \begin{pmatrix} X \\ \xi \end{pmatrix} = b + \tilde{B} \begin{pmatrix} e \\ e \\ \vdots \\ e \end{pmatrix} \quad \text{for } e > 0$$

$$= b + e \begin{pmatrix} r_1 \\ \vdots \\ r_M \end{pmatrix}$$

where r_i is the sum of the i^{th} row of \tilde{B} .

Then

$$79) \quad \begin{pmatrix} X_j(e) \\ \xi_i(e) \\ \xi_{\bar{B}} \end{pmatrix} = \tilde{B}^{-1} b + \begin{pmatrix} e \\ \vdots \\ e \end{pmatrix}$$

Thus the original optimal basis is still feasible and therefore uniquely optimal (since it still satisfies the pricing relationships).

Also for $\epsilon > 0$

$$x_j(\epsilon) > 0 \quad \text{for } j \in \mathcal{J}$$

$$\xi_i(\epsilon) > 0 \quad \text{for } i \in \bar{\mathcal{I}}$$

and

$$\begin{pmatrix} x_j(\epsilon) \\ \xi_i(\epsilon) \end{pmatrix} \longrightarrow \begin{pmatrix} x_j^0 \\ \xi_i^0 \end{pmatrix}$$

as $\epsilon \rightarrow 0$

The procedures of the last section which apply when $L_{\mathcal{B}}$ has a unique and non-degenerate solution apply with essentially no modification if $L_{\mathcal{B}}$ has a unique but possibly degenerate solution if we let $P(\epsilon)$ be

$$\min V = \sum a_{ij} x_i x_j + \sum e^{j+1} x_j$$

subject to

$$80) \quad \sum a_{ij} x_j = b_i + r_i \epsilon + e^{N+i+1} \quad \text{for } i = 1, \dots, m_1$$

$$\sum a_{ij} x_j \geq b_i + r_i \epsilon + e^{N+i+1} \quad \text{for } i = m_1 + 1, \dots, n$$

The solution to $L_{\mathcal{B}}(\epsilon)$ is non-degenerate for sufficiently small ϵ . Along any critical line we now have

$$81) \quad x_{j_B} = \alpha_{xj_B} + \beta_{xj_B} \lambda_B + \epsilon P_{xj_B}(\epsilon) + \left(\sum_{h=1}^I m^{B,h+j} r_{1h} \right) \epsilon$$

$$= \alpha_{xj_B} + \beta_{xj_B} \lambda_B + \epsilon Q_{xj_B}(\epsilon)$$

$$82) \quad \lambda_{i_B} = \alpha_{\lambda i_B} + \beta_{\lambda i_B} \lambda_B + \epsilon P_{\lambda i_B}(\epsilon) - \left(\sum_{h=1}^I m^{B+j,h+j} r_{1h} \right) \epsilon$$

$$= \alpha_{\lambda i_B} + \beta_{\lambda i_B} \lambda_B + \epsilon Q_{\lambda i_B}(\epsilon)$$

$$83) \quad \xi_1 = \alpha_{\xi_1} + \beta_{\xi_1} \lambda_B + \sum_{s=1}^J a_{1j_s} a_{2j_s}(e) + e^{B+i+1}$$

$$= \alpha_{\xi_1} + \beta_{\xi_1} \lambda_E + a_{\xi_1}(e)$$

$$84) \quad \eta_j = \alpha_{\eta_j} + \beta_{\eta_j} \lambda_E + \sum_{s=1}^J a_{jj_s} a_{2j_s}(e) - \sum_{s=1}^I a_{1s_j} a_{2s_j}(e) + e^{j+1}$$

$$= \alpha_{\eta_j} + \beta_{\eta_j} \lambda_E + a_{\eta_j}(e)$$

where the $p_v(e)$ are as defined in 63) through 66). Since no $p_v(e)$ can have zero coefficients and no two can have proportional coefficients, the same is true of the $a_v(e)$.

9. The Algorithm when L_B is not Unique

A non-degenerate optimal solution to $L_{\bar{B}}$ is unique if and only if

$$85) \quad \delta_j < 0 \text{ for } X_j \text{ not in the basis}$$

$$p_i > 0 \text{ for } \xi_i \text{ not in the basis.}$$

If L_B has a degenerate solution and 85) does not hold, then either the solution is not unique or else only the optimal basis is not unique. If L_B does not have a unique solution we must find the point \bar{X} which gives minimum V for $Z = \bar{E} = \max E$. If only the optimal basis of $L_{\bar{B}}$ is not unique we still must decide on the \mathcal{J} of our first critical line. Both these problems will be resolved in the same manner. Our procedure may be considered as a special case of either approach 1 or approach 4 for minimizing a quadratic subject to linear constraints described in section 12.

Let us create a new linear programming problem $L_p(e)$ by adding a constraint to and modifying the form to be maximized in $L_p(c)$. The equation we add is

$$86) \quad \sum_{j=1}^N \mu_j X_j - \xi_E = \bar{E} + \left(\sum_{j \in J} \mu_j - 1 \right) e$$

If we add ξ_E to the optimum basis variables of $L_p(c)$ we have a feasible basis corresponding to a solution with

$$87) \quad \begin{aligned} X_j &= X_j^0 + e & j \in J \\ \xi_i &= \xi_i^0 + e & i \in \bar{J} \\ \xi_E &= e \end{aligned}$$

Next let us replace the objective function $E = \sum \mu_i X_i$ with a new one

$$88) \quad F = \sum v_j X_j$$

such that the solution in 87) is the unique optimum of $L_p(e)$. This may be done easily by assigning any values $p_i > 0$ to $i \notin \bar{J}$, $p_i = 0$ for $i \in \bar{J}$ as well as $p_E = 0$. Then choose any set of v_j so that $\delta_j = 0$ for $j \in J$ and $\delta_j < 0$ for $j \notin J$. Since $L_p(e)$ has a unique non-degenerate solution we may use methods already described to trace out the set of points which give minimum $V(e)$ for given F until $\lambda_p = 0$. If only a few bases are feasible for $L_p(c)$, i.e., if not too many bases are optimal for $L_p(e)$, $\lambda_p = 0$ will be reached quickly. At $\lambda_p = 0$ either $\lambda_E = 0$ or $\lambda_E > 0$. In the former case we have arrived at a point with minimum V and maximum E . In the latter case we have \bar{X} and are ready to trace out the set of efficient X 's. From this point on we let $\lambda_p = 0$, i.e., we ignore F completely. Since at $\lambda_p = 0$, $v_k > 0$ for

all $k \neq X$ we may reduce $\lambda_{\underline{X}}$ until we intersect a plane $v_k = 0$ and continue as in section 8.

10. The Algorithm, when Condition 3 does not Hold

If E is unbounded procedure 4 of section 12 can be used to find the point \underline{X} with minimum V . The efficient set can then be traced out in the direction of increasing $\lambda_{\underline{X}}$. Since there are only a finite number of critical lines and each critical line is efficient at most once the efficient set is traced out in a finite number of steps.

11. The Set of Efficient E,V Combinations

Once the set of efficient X 's is found the set of efficient E, V combinations can be obtained easily. The critical line of a subspace in which more than one value of E is obtainable may be expressed as the solution to

$$89) \quad \sum \sigma_{jk} x_k + \sum (-\lambda_i) a_{ij} + (-\lambda_E) \mu_j = 0, \quad j \in \mathcal{J}$$

$$90) \quad \sum a_{ij} x_j = b_i, \quad i \in \mathcal{I}$$

$$91) \quad \sum \mu_j x_j = E$$

for $-\infty < E < +\infty$

If we let N^{-1} be the inverse of the matrix N in 80), 90), 91) we have

$$92) \quad \begin{pmatrix} \underline{X} \\ -\lambda \end{pmatrix} = N^{-1} \begin{pmatrix} 0 \\ b \\ E \end{pmatrix}$$

$$93) \quad V = (X', -\lambda') \begin{pmatrix} C & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \underline{X} \\ -\lambda \end{pmatrix} = (0', b', E) N^{-1} C N^{-1} \begin{pmatrix} 0 \\ b \\ E \end{pmatrix}$$

from which it follows that along any such critical line V and E are related by a formula of the form

$$94) \quad V = a + bE + CE^2.$$

Thus the set of efficient E,V combinations is piecewise parabolic. We know, or can easily get, the values of E and $\frac{dV}{dE}$ at the end points of each of the pieces. We can also evaluate V at \bar{X} .^{*} Knowing V at one value of E and $\frac{dV}{dE} = b + 2CE$ at two values of E we can solve for the a, b, and c in 94) for the segment from \bar{E} to $\bar{E} - e_1$. Having a, b, and c we can evaluate V at $\bar{E} - e_1$ by means of 94). This provides us with the value of V at one value of E on the segment which is efficient from $\bar{E} - e_1$ to $\bar{E} - e_2$. This, combined with the values of $\frac{dV}{dE}$ at two values of E, allows us to obtain the a, b, c of 94) for this next segment--and so on until we trace out the set of E,V combinations.

12. Minimizing a Quadratic

One of the "by-products" of the calculation of efficient sets is the point at which V is a minimum, i.e., where $\lambda_{\bar{X}} = 0$. The computing procedures described in sections 6 through 10 are analogous to the simplex method of linear programming (as contrasted with the "gradient methods" that have been suggested for both linear and non-linear programming). Both the procedure described in the preceding section--considered as a way of getting to min V--and the simplex method require a finite number of iterations, each iteration typically taking a "jump" to a new point which is superior to the old. Each iteration makes use of the inverse of a matrix which is a "slight" modification of the matrix of the previous iteration. The success of the simplex method in linear programming suggests that it may be desirable to use a variant of the "critical line" method in the quadratic case.

^{*}If \bar{X} does not exist we can evaluate V at \underline{X} and use the same process "in reverse."

Our problem then is to minimize a quadratic

$$V = \sum a_{ij} X_i X_j$$

subject to constraints 1), 2) and 3). We wish to translate this into a problem of tracing out an efficient set. This may be done in several ways.

1. An arbitrary set of μ_j can be selected and the efficient set traced out until $\lambda_E = 0$. The μ_j should be selected so that the "artificial" E has a unique maximum.

2. An equality, say,

$$\sum a_{1j} X_j = b_1$$

can be eliminated from 1). E can be defined as

$$E = \sum a_{1j} X_j$$

and the critical set traced out until $E = b_1$. If $E = b_1$ is reached before $\lambda_E = 0$ the computing procedures of the last section must be continued into the region of $\lambda_E < 0$. While the points thus generated will not be efficient--for they do not give $\max E$ for given V--they do give $\min V$ for given E. In particular they will arrive at the point of $\min V$ for

$$E = \sum a_{1j} X_j = b_1$$

3. An inequality, say,

$$\sum a_{mj} X_j \geq b_m$$

can be eliminated from 2). E can be defined as

$$E = \sum a_{mj} X_j$$

The efficient set is traced out until either $E = b_m$ or else $\lambda_E = 0$.

If the former happens first the constraint is effective; if the latter

happens first the constraint is ineffective. In either case the point associated with the first of these to occur gives min V subject to 1), 2), and 3).

4. An initial guess X_1^0, \dots, X_N^0 which satisfies 1), 2), and 3) can be made and μ_j defined so that, given these μ_j , X^0 is efficient. The efficient set can then be traced out until $\lambda_{X^0} = 0$. To choose μ_j so that X^0 is efficient choose arbitrary positive values of λ_1 (i.e. 1) and λ_N . Then choose μ_j so that

$$\eta_j = 0 \quad \text{for } X_j \text{ not at its lower bound}$$

$$\eta_j > 0 \quad \text{for } X_j \text{ at its lower bound}$$

If X^0 is in the same subspace as the optimal solution, the latter is reached in one iteration.